Addressing the Ethical Implications of Generative Artificial Intelligence in K-12 Education Through Innovation: Developing an Intelligent Academic Assistant to Support Critical Thinking

Jesus Barrios
Amityville Memorial High School
Amityville, NY
jesbarri@s.amityvilleufsd.org

Cristian Reyes
Amityville Memorial High School
Amityville, NY
crireyes@s.amityvilleufsd.org

Word Count: 5,297

1 Introduction

Have you ever used artificial intelligence (AI) to help you complete a task? The usage of AI has expanded rapidly. AI refers to programs and algorithms designed to perform processes that mimic human intelligence, such as understanding language, identifying patterns, and solving problems. This field contains many different subfields, such as machine learning and natural language processing. This paper will focus on generative artificial intelligence (GenAI), which is AI designed to create new content, such as text and images.

The area of focus for this project will be in K-12 education, specifically adolescents. This paper will explain the history of AI, its implications, ethical and social concerns, and a proposed solution for K-12 schools, a critical stage of adolescent development where GenAI use can have significant implications. This proposed solution presented at the Long Island Youth Summit is the development of an Intelligent Academic Assistant, an AI-powered full-stack application that allows students to input academic queries and receive helpful outputs. This model will address important issues explained later in this paper by serving as an assistant, supporting critical thinking in its responses.

The current situation of GenAI use in K-12 schools across Long Island, New York, is an outright ban. School districts accomplish this through network and chromebook settings that block student access. The proposed solution of an Intelligent Academic Assistant challenges these complete bans by addressing the ethical concerns surrounding GenAI use. Although this solution will address GenAI use within schools, the question still remains about usage outside of the classroom since these areas cannot be regulated by the district. Would legislation work? How effective would a ban be? These issues will be discussed at the end of the paper following the presentation of the Intelligent Academic Assistant.

The primary purposes of this project are as follows:

- Firstly, to develop a prototype Generative AI academic assistant that supports critical thinking and learning while mitigating academic dishonesty. This prototype will serve as an alternative to the outright ban of Generative AI in schools.
- Secondly, to initiate dialogue among Long Island school districts regarding the ethical integration of AI, aligning with the perspectives of experts who advocate for AI fluency among students.
- Thirdly, to provide a tangible demonstration of a Generative AI resource that supports student learning and cognitive development, thereby encouraging schools to research and implement similar solutions.

This solution is a clear demonstration of GenAI's potential to support education when used responsibly. The desired purpose of the project is to create an academic assistant that supports critical thinking and ethical use in its responses, which will be achieved during its development phase. GenAI, when used responsibly, can be a beneficial

resource that isn't used for academic dishonesty. This approach aligns with the opinions of various education professionals in identifying that AI is the future and AI fluency needs to be taught. Ultimately, this project serves as an invitation to promote AI-supported learning and encourage schools to allow this type of usage from alternative models that hold educational values.

2 The History of Artificial Intelligence

The beginning of AI can be traced back to the 1950s at Dartmouth College. A group of researchers gathered to investigate the possibility of machines that perform tasks similar to humans, such as using language, forming abstractions, and solving problems. This meeting laid the foundation for AI's development over the next few decades. In fact, this is where John McCarthy coined the term "artificial intelligence." Early research focused on programs that combined facts with logic through probabilistic models. By the 1980s, though, symbolic logic became the favored approach. However, early versions of AI were proving unreliable and, as a result, became less popular due to overpromising and underdelivering [12]. Although early development of AI was extremely ambitious, there were significant challenges, like lack of infrastructure, that limited its potential.

Nevertheless, advancements in computer hardware eventually created a path forward. In 1986, a backpropagation algorithm was developed by Rumelhart, Hinton, and Williams, which allows neural networks to learn complex patterns. More progress occurred in 1988, when Judea Pearl's work on Bayesian networks allowed AI systems to make decisions about uncertain situations. Furthermore, in 1997, Long Short-Term Memory networks (LSTMs) were created by Hochreiter and Schmidhuber, which overcame the earlier neural network limitations in capturing long-range relationships in data. LSTMs are widely used in handwriting and speech recognition, language processing, and more. As a result, all of these advancements paved the way for the development of OpenAI's GPT-3 in 2020, which swept the market on release [13]. GPT-3 showed a remarkable ability to generate human text and maintain a conversation, showing world-changing potential in the future of GenAI. Significant progress had been made since the first meeting at Dartmouth College, and OpenAI's release of GPT-3 popularized the field.

Looking forward, the future of AI is increasingly ambitious. Project Stargate was recently announced, which is a project involving major tech companies such as OpenAI, SoftBank, Oracle, and Nvidia. This is a $500 billion investment in United States AI infrastructure over the next four years. The goal is to outcompete competition in foreign countries like China and to support the development of artificial general intelligence (AGI), an AI with human-level intelligence [6]. This project depicts that AI is here to stay and is worth investing in.

3 GenAI Usage

Since the release of GPT-3, the usage of GenAI has increased rapidly, becoming widespread across the world. Several competitors have risen in an attempt to lead the AI boom, such as Gemini by Google, Claude by Anthropic, and recently Deepseek. Salesforce's [2023] Generative AI Snapshot Research Series surveyed over 4,000 individuals on GenAI use in various different categories. The data was converted into a table depicted below.

Table 1: Generative Artificial Intelligence Usage Across Demographics and Professions (n>4000)

Category	Subcategory	Percentage
Geographical Usage	India	73%
	Australia	49%
	United States	45%
	United Kingdom	29%
Generational Trends	Millennials & Gen Z (Users)	65%
	Gen Z Using AI	70%
	Non-Users: Gen X & Baby Boomers	68%
Employment & AI Use	Employed AI Users	72%
	AI Users Automating Work Tasks	75%
	AI Users Increasing Usage Over Time	52%
Attitudes Toward	Users Who	67%

AI	Believe AI Benefits Work	
	Users Who Believe AI is Risky	73%
Workplace Adoption	Workers Currently Using AI	61%
	IT Leaders Prioritizing AI	67%
	Marketers Using AI	51%
	Sales Professionals Using AI	33%
	Customer Service Professionals	24%

Note: Data sourced from Salesforce's Top Generative AI Statistics for 2025 [10].

Table 1 illustrates the widespread usage of GenAI. India has the highest usage (73%), followed by Australia (49%), the United States (45%), and the United Kingdom (29%). Additionally, Millennials & Gen Z are the most active users with 65% usage, with 68% of Gen Z and Baby Boomers remaining non-users.

Taking a look at GenAI usage in the workplace, 72% of employed users are using GenAI, with 75% automating work tasks, with a 52% increase in usage over time. Even though 73% of users believe GenAI use is risky, there's still a wide amount of usage. In specific professions, workers are reporting 61% usage, IT leaders 67%, marketers 51%, sales professionals 33%, and customer service professionals 24%.

This data shows the high usage of GenAI across various fields. Many individuals have begun using GenAI for assistance in the workplace and for general purposes.

3.1 GenAI Usage in K-12 Education

How do these statistics compare to those of K–12 students? The following table was created from data reported by ACT [2023], based on a sample of 4,006 students in grades 10–12.

Table 2: High School Students' GenAI Usage and Attitudes (n=4006)

Category	**Subcategory**	**Percentage**
Overall Usage	Students reporting using AI tools	46%
	Using AI for school assignments *(Among users)*	46%
Reasons for Not Using AI Tools	Lack of interest	83%
	Not trusting the information provided	64%
	Not knowing enough about AI tools	55%
Usage by Subject	Language Arts	66%
	Social Studies	49%
	Science	37%
Teacher Restrictions	Teachers not allowing AI tool use for schoolwork	62%
Common AI Tools Used	ChatGPT	83%
	Dall-E-2	17%
	Bing Chat	11%
Accuracy Concerns	Students reporting errors/inaccuracies in AI-generated responses	63%
Academic Achievement	Top quartile ACT scores using AI tools	53%
	Bottom quartile ACT scores using AI tools	36%
Opinions on Banning AI Tools	Students supporting a ban on AI tools	42%
	Students opposing a ban on AI tools	34%
	Students unsure about banning AI tools	23%
College Admissions Essay	Students considering using AI for writing college admissions essays	10%

Note: Data adapted from ACT's report [1].

Table 1 illustrates the significant usage and opinions of GenAI across high school students. 46% of students have reported GenAI use, with 46% of those users using it for school assignments. Among the non-users, 83% reported a lack of interest, 64% reported distrust, and 55% reported not understanding much about it. It seems like Language Arts (66%) and Social Studies (49%) are the highest subjects of GenAI usage, suggesting GenAI is being used to complete writing assignments. ChatGPT was reported as the most popular GenAI tool at 83%. Teachers have begun addressing this, with 62% prohibiting its usage for classwork. Additionally, GenAI isn't only being used by low-achieving students. There is significant usage (53%) among students with top ACT scores. This reinforces the widespread usage of GenAI. In fact, GenAI has become so popular that there is no significant majority of students that support its ban.

4 Rationale for Focusing on K-12 Education

The significant usage of GenAI in high schools seen in Table 2 is a significant reason for the focus of this study on K-12 education. This data was released in 2023, and it's reasonable to assume that GenAI usage has only increased due to growing popularity and easy access. GenAI usage and its implications in groups outside of adolescents, which is significant and widespread, are outside the scope of this project. The focus on adolescents narrows the scope of the study, allowing more focused and meaningful analysis.

4.1 Adolescent Development

The primary reason for a focus on adolescents (ages 10–19) is due to the fact that it's a critical stage of development that has a significant impact. Backes and Bonnie [2019] explain that "adaptive plasticity" occurs at this stage, which is an opportunity for positive growth and change due to the brain's malleability, adapting in response to its environment. However, this plasticity also makes adolescents susceptible to negative influences, such as drugs or violence. This can "get under the skin" and affect development. Although GenAI usage is nowhere near comparable to drug usage, these findings suggest a critical stage of development where adolescents are becoming independent and self-sufficient in response to the environment. As a result, GenAI usage may hinder this process and raise concerns on the development of critical thinking.

5 Impact of GenAI on K-12 Education

5.1 GenAI and Academic Dishonesty

This finding is of particular concern when considering the significant usage of GenAI for academic dishonesty, which is the process of gaining an unfair advantage in an academic setting. Turnitin [2024] reported that its AI detector identified AI-generated content in about 10% of 200 million writing assignments reviewed, with 3% being entirely written by AI. This is becoming a significant concern for educators, with a survey by the Center for Democracy & Technology [2024] reporting that 68% of middle and high school teachers are implementing AI detectors during the 2023-24 school year. Of these teachers, 63% (up from 48% the year prior) reported disciplining students for AI use, which reinforces Turnitin's results of widespread academic dishonesty. Half of the teachers surveyed mentioned an increased distrust in student work, raising concerns of authenticity. These findings support the widespread usage of GenAI for academic dishonesty, which raises concerns about its long-term impact on the development of critical thinking.

5.2 Effects of GenAI on Critical Thinking Skills

The concerns introduced from the work of Backes and Bonnie [2019], Turnitin [2024], and the Center for Democracy & Technology [2024] on critical thinking are confirmed in new studies. Lee et al. [2025] found a positive relationship between high confidence in GenAI and reduced critical thinking, suggesting that reliance on these resources weakens problem solving. Additionally, GenAI seems to lessen the perceived effort required for critical thinking. Lee et al. [2025] also notice "mechanized convergence," where work assisted by GenAI leads to less diverse outcomes. These findings suggest that GenAI usage may reduce critical thinking, which is particularly worrisome when considering adolescents are using this resource during a sensitive critical stage of development.

This raises questions on how to address GenAI usage in K-12 schools. Should it be banned? The experts are saying not. It seems like the concern arises due to the fact that GenAI is directly providing answers and doing tasks that the students should be completing. Spector [2023], in an interview with Stanford education professionals, states that cheating has always been an issue, with surveys over the past 15 years showing 60-70% of students cheat in some form each month. Stanford professionals acknowledge the significant potential for

supporting learning. Instead of banning AI chatbots, educators should address the causes of cheating and teach responsible AI use. This approach shows a new push in teaching AI fluency, considering its potential to teach students. Many students agree that AI is acceptable for introductory purposes, such as explaining new concepts. This concept will be the basis of the proposed solution introduced later. However, before that, it's important to address the ethical and social implications of GenAI usage.

6 Ethical and Social Considerations of GenAI in Education

The three main ethical concerns surrounding GenAI in K–12 education are privacy, bias, and, most importantly, the role of human judgment.

6.1 Privacy Risks Associated with GenAI Use

The privacy concerns associated with GenAI usage are concerning. O'Neill [2025] explains that GenAI resources risk potential breaches that could release sensitive user information, such as conversation history. These risks can be reduced through methods like data anonymization, which removes personal identification before training GenAI models. Additionally, allowing data processing and machine learning to occur on the local device could reduce the risks posed with transferring data. Privacy is extremely important and a right that many individuals expect. Therefore, GenAI models must be held to significant security standards to ensure user data is protected.

6.2 Bias in GenAI

In addition to privacy, bias is another significant ethical concern. In fact, all GenAI models are inherently biased due to the data used to train them. Datta [2025] explains "native legacy bias," which refers to pre-existing biases in the data used to train models. For instance, language translation models have associated female names with terms like "parents" and "weddings." On the other hand, male names were connected to words like "professional" and "salary." The data used to train these models unintentionally reflects the existing biases in society and the years of systemic bias against certain groups. This bias trained into GenAI models reinforces harmful stereotypes and exposes them to adolescents using it, which may unintentionally leave an impact.

6.3 The Diminishing Role of Human Judgment

Finally, the most significant concern regarding GenAI usage is the role of human judgment. The work of Lee et al. [2025], as discussed earlier, depicts how GenAI can reduce critical thinking. This raises the question: What is the purpose of human judgment when an AI can make decisions for you? GenAI has the potential to reduce the effectiveness of human judgment and possibly reduce its need. Gerlich's [2025] study shows that although AI use can free up cognitive resources for more complex thought, it can also decrease the need for deep cognitive engagement, which leads to a decline in memory and critical analysis. This suggests that as individuals become more dependent on AI, their capacity for critical thinking and reflective reasoning may diminish. This poses serious ethical and social concerns, specifically on adolescents, since GenAI is diminishing thought.

7 Methodology and Solution

7.1 Action Research: Design and Approach

The methodological approach for this project will be action research, which involves both investigating an issue and attempting to solve it. Based on the ethical and social implications of GenAI usage in K-12 education, the next step is to address them through a proposed solution. This approach is valid because it lays a foundation for future research that builds upon this project.

The proposed solution for the Long Island Youth Summit's Science and Technology topic is the development of an Intelligent Academic Assistant, an AI-powered academic helper. This solution was created to address the ethical concerns of GenAI use, specifically the role of human judgment and the rise of academic dishonesty, by creating a GenAI model that supports critical thinking. Its development will serve three primary purposes, as introduced in the introduction:

- Firstly, to develop a prototype Generative AI academic assistant that supports critical thinking and learning while mitigating academic dishonesty. This prototype will serve as an alternative to the outright ban of Generative AI in schools.
- Secondly, to initiate dialogue among Long Island school districts regarding the ethical integration of AI, aligning with the perspectives of experts who advocate for AI fluency among students.

- Thirdly, to provide a tangible demonstration of a Generative AI resource that supports student learning and cognitive development, thereby encouraging schools to research and implement similar solutions.

This solution is building upon the thoughts of Spector [2023], who states that GenAI must be incorporated responsibly. This is supported by the work of Nagelhout [2024], whose study shows that many participants have described GenAI as "the modern approach to learning," acknowledging that "not all kids use it to cheat in school." The Intelligent Academic Assistant is designed to follow up on these ideas and address the concerns related to GenAI use.

7.2 System Architecture of the Intelligent Academic Assistant

The Intelligent Academic Assistant will be a trained model powered by OpenAI's gpt-3.5-turbo-0125 model. This choice was made due to the model's cost efficiency and reliability. The creation of a fresh large language model was ruled out of the picture due to a lack of resources and a high likelihood of underperformance compared to existing models. The gpt-3.5-turbo-1025 model will be fine-tuned with data on desired conversational outcomes. Fine-tuning is the process of training a pre-trained model on specific data, which means training the model on datasets of helpful conversations, educational content, and responses that support critical thinking. The goal is to create an Intelligent Academic Assistant that guides students through the learning process and supports critical thinking.

The goal isn't to only just create a model, but also to create a frontend to allow testing and previewing of its capabilities. This prototype will allow users to create accounts and manage multiple conversations. This data will be stored within a PostgreSQL database, using Prisma's Object-Relational Mapper to interface it. All data, such as account information and conversations, will be safely stored. It's important to note that Google OAuth will be used for user login, which removes the risk with password management, and that the data is protected through Transparent Data Encryption, which prevents unauthorized access. All conversation data, including inputs and outputs, will be saved to the database to maintain contextual awareness through user inputs. Whenever a user submits a prompt, a chronological array of conversation history will be submitted to the model, which ensures that the model remembers previous queries. This is designed to handle large amounts of data without reducing performance.

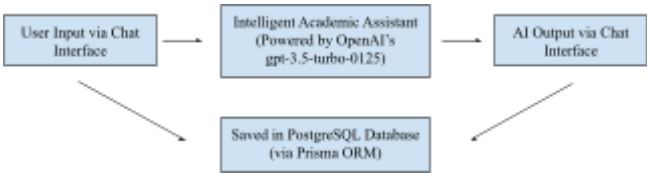
User Input via Chat Interface
Intelligent Academic Assistant (Powered by OpenAI's gpt-3.5-turbo-0125)
AI Output via Chat Interface
Saved in PostgreSQL Database (via Prisma ORM)

Figure 1: Simplified Process Flow of the Intelligent Academic Assistant

Figure 1 depicts a simplified representation of the overall process flow of the Intelligent Academic Assistant. The user submits an input, which is processed through the Intelligent Academic Assistant (the fine-tuned model) and produces an output. It also depicts the saving of data to maintain memory throughout the entire conversation.

7.3 Detailed Implementation Strategy

7.3.1 Fine-tuning the GPT-3.5-Turbo Model

The fine-tuning of the gpt-3.5-turbo-0125 model was done within OpenAI. Supervised learning was chosen as the best method to prepare the model with the patterns needed for an academic assistant. The data to train the model was prepared inside of a JSON line file.

```
{
  "messages": [
    {
      "role": "system",
      "content": "You are an Intelligent
      Academic Assistant. You are not
      allowed to give the direct answer
      to a request. Instead, you must
      serve as a tutor for students and
      guide the student to the right
      direction."
    },
    {
      "role": "user",
      "content": "Find the slope of the
      line passing through the points
      (2,3) and (5,11)."
    },
    {
      "role": "assistant",
      "content": "To find the slope
      between two points, use the slope
      formula: m = (y2 - y1) / (x2 - x1).
      Can you plug in the values from the
      points (2,3) and (5,11)?"
    }
  ]
}
```

Figure 2: Sample JSON Data for Training the Model

Figure 2 depicts a sample of data inside of the JSON line file used to train the model. This is just one of 200 conversations used to train the model, which can be found in Appendix A.1. This data was created to cover the four core academic subjects: English, Social Studies, Math, and Science. Each subject had 50 conversations each that represented it. It's important to note that the pre-existing capabilities of the gpt-3.5-turbo-0125 model are extremely high, which means that this assistant will be able to cover a wide range of different subjects. The training data was designed to simply solidify the model's role as an assistant and depict desired outcomes. It's also important to note that 200 conversations for training is an extremely small amount, which may impact the results received from the model. 200 conversations were chosen due to the lengthy process of creating datasets, which may affect the results received from the model. This should be expanded in future studies.

After uploading the data, a random seed was chosen to ensure reproducibility, and both the batch size and learning rate multiplier were set to 1. The model was trained over 3 epochs, which means that the model processes the entire dataset 3 times, which is appropriate for this prototype phase. All of these parameters play an important role in controlling the speed and accuracy of the training process. Finally, the job was submitted, and the Intelligent Academic Assistant model was created.

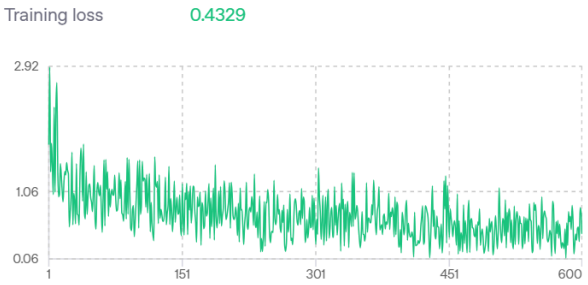

Figure 3: Training Loss Graph during Fine-Tuning of Model

Figure 3 shows a training loss graph that depicts the model's learning trajectory across the three epochs. This graph shows how the model's error decreased over time, beginning with a high value and decreasing to a small 0.4329 training value. This relatively low value suggests that the model effectively learned the patterns in the training data, confirming a successful fine-tuning process.

7.3.2 Implementing the Intelligent Academic Assistant

```
Input: API Key, Message, Optional List of Previous Messages
Output: AI Chat Completion Choice

Create a list of all messages:
  allMessages = [ { role: "user", content:
  process.env.PROMPT } ] + messages + [ { role: "user",
  content: message } ]

Create a chat completion request:
  chatCompletion = ai.chat.completions.create
    model = "ft:gpt-3.5-turbo-0125:jesus-development:
    intelligentacademicassistant:B2RnNRsR"
    max_tokens = 1000
    temperature = 0
    messages = allMessages

Return chatCompletion.choices[0]
```

Figure 4: Incorporation of Fine-Tuned Model for AI Chat Completion

Following the successful fine-tuning of the gpt-3.5-turbo-0126 model to create the Intelligent Academic Assistant model, the next step was incorporating it into the backend of the project. Figure 4 depicts a representation of its implementation in the backend, specifically by using OpenAI's API to create a completion. create call, which generates a chat completion based on the prompt. Figure 4 also depicts how the model retains conversational context, or memory, when creating a completion. An array containing all conversation history, retrieved from the PostgreSQL database, was submitted as input, allowing the model to maintain context. This was done in TypeScript, and the entire source code can be found in Appendix A.2.

> Prompt: You are an Intelligent Academic Assistant. You are not allowed to give the direct answer to a request. Instead, you must serve as a tutor for students and guide them to the right direction. You must provide information that will lead the student to find the answer on their own.

Figure 5: Prompt for Chat Completion

The prompt referenced in Figure 4 and shown in Figure 5 was also submitted into the model as the first item in the conversation. The purpose was to match the prompt and training data used to fine-tune the model in an effort to improve performance and achieve our desired outcome of

operating as an assistant. The prompt tells the model to guide students and to support a learning environment that promotes critical thinking.

After the implementation of the model into the backend, the next step was creating a frontend interface that allows testing of its capabilities. The interface was created inside of TypeScript using React syntax. The frontend allows users to manage their accounts, manage conversations, and chat with the trained model. This completed the development of the Intelligent Academic Assistant. The training data used for fine-tuning can be seen in Appendix A.1, and the complete open-source code repository is available in Appendix A.2. The project can be accessed at https://intelligent-academic-assistant.vercel.app/.

8 Results

The first metric measured was response time. The response time of the Intelligent Academic Assistant was tested using prompts designed to replicate the levels of complexity a student might ask. Each prompt was executed 10 times in fresh conversations, and the average response time for each was recorded.

Table 3: Query Complexity and Average Response Time

Query Complexity	Query	Average Response Time (s)
001	What is the capital of France?	1.58
002	Define photosynthesis.	1.91
003	What is 123 multiplied by 456?	1.99
004	Explain the concept of supply and demand in economics.	2.10
005	Compare and contrast the American and French Revolutions.	2.28
006	Write a Python function to calculate the factorial of a number.	2.21
007	Write a short story about a robot who discovers emotions.	2.38
008	Explain the theory of relativity and its implications for space travel, in a way a high school student could understand.	3.03

Table 3 depicts the results of the response time evaluation. When asked simple fact-based questions such as "What is the capital of France?", the model responded in 1.58 seconds. This time increased as query complexity increased. For instance, when asked complex prompts such as “"Explain the theory of relativity and its implications for space travel in a way a high school student could understand," the model responded in 3.03 seconds. This suggests that the model can handle all types of questions, powered by gpt-3.5-turbo, though the timing tends to increase as complexity increases. Although the model likely took milliseconds to process and produce the information, it took a while to display it on our frontend. This might’ve been influenced by factors such as internal processing overhead, network latency, and environmental limitations.

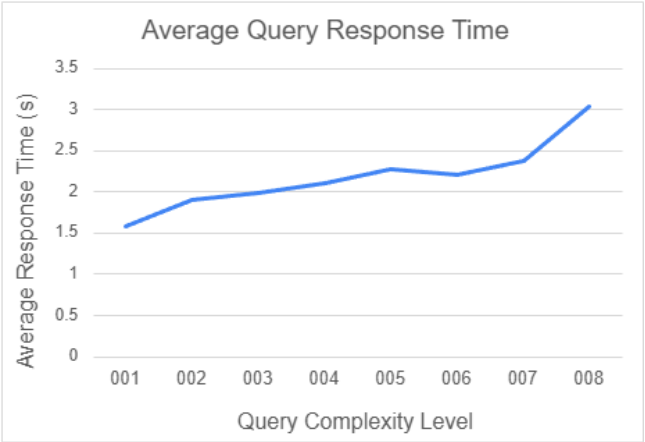

Figure 6: Relationship Between Query Complexity and Average Response Time

This trend can be better visualized in Figure 6. While the performance is acceptable given our resources, further studies must optimize their developmental environment to reduce these delays and improve the user experience.

The second metric examined was the most important one: the effectiveness of the Intelligent Academic Assistant. A qualitative analysis was performed to assess the quality of the responses to ensure they align with our desired purpose of supporting critical thinking. The model

was tested using 20 conversations, with 5 prompts for each of the four core subjects. This can be seen in Appendix A.3.

Figure 7 below depicts the criteria used to examine the effectiveness of the responses received in the 20 conversations. Each conversation will receive 7 ratings of low, medium, or high, each aligning with one of the criteria.

Criteria	**Descriptions**
Clarity	How easily can the user understand the response?
Context Retention	Does the response preserve all the necessary background from the original prompt?
Helpfulness	Does the response offer genuine advice?
Accuracy	Is the information correct and does it align with the task's requirements?
Engagement	Does the response encourage further thought and interaction?
Supports Critical Thinking	Does the response include prompts or questions that support deeper analysis?
Avoids Direct Answers	Does the response guide the user rather than simply providing a complete answer?

Figure 7: Assistant Response Analysis Criteria and Descriptions

ID	Clarity	Context Retention	Helpfulness	Accuracy	Engagement	Supports Critical Thinking	Avoids Direct Answers
001	High	High	High	High	High	High	High
002	High	High	High	High	High	High	High
003	High	High	High	High	High	High	Low
004	High	High	High	High	High	High	High
005	High	High	High	High	High	High	High
006	High	High	Medium	High	High	High	High
007	High	High	High	High	High	High	High
008	High	High	High	High	High	High	High
009	High	High	High	High	High	High	Low
010	High	High	High	High	High	High	Low
011	High	High	High	High	High	High	High
012	High	High	Medium	High	High	High	High
013	High	High	High	High	High	High	High
014	High	High	High	High	High	High	High
015	High	High	Medium	High	High	High	High

016	High	High	High	High	High	High	Low
017	High	High	High	High	High	High	Low
018	High	High	High	High	High	High	High
019	High	High	High	High	High	High	Low
020	High	High	High	High	High	High	Low

Figure 8: Assistant Response Performance Based on Criteria

Figure 8 depicts the results of the responses (seen in Appendix A.3) based on the criteria shown in Figure 7. Overall, the Intelligent Academic Assistant performed very well, mostly due to its power from the gpt-3.5-turbo-0125 model. The majority of responses were rated high in terms of clarity, context retention, accuracy, engagement, and support for critical thinking. There were some discrepancies, though, which will be addressed.

Table 4: Frequency Table of Criteria Ratings

Criteria	High (Count / %)	Medium (Count / %)	Low (Count / %)
Clarity	20 (100%)	0 (0%)	0 (0%)
Context Retention	20 (100%)	0 (0%)	0 (0%)
Helpfulness	17 (85%)	3 (15%)	0 (0%)
Accuracy	20 (100%)	0 (0%)	0 (0%)
Engagement	20 (100%)	0 (0%)	0 (0%)
Supports Critical Thinking	20 (100%)	0 (0%)	0 (0%)
Avoids Direct Answers	13 (65%)	0 (0%)	7 (35%)

Table 4 depicts a table used to organize the findings seen in Figure 8. The model received perfect high scores in all 20 conversations in clarity, context retention, accuracy, engagement, and support of critical thinking. However, there were setbacks in helpfulness and avoiding direct answers.

35% of the 20 conversations received a low rating for avoiding direct answers. This means that in the 7 conversations, the assistant provided more direct answers than intended, which might've reduced the opportunity to engage in problem solving, since the model did it for them. However, the model was simply attempting to be helpful and responsive and might not have intended to offer a complete answer for cheating.

For instance, in conversation 003, the prompt was "Explain the difference between dramatic irony, situational irony, and verbal irony." The assistant responded by providing a direct explanation of each term, causing a low rating for avoiding direct answers since it did not encourage the student to figure it out on their own. This raises a challenging question: How can the system distinguish between genuinely helping a student understand a concept and unintentionally allowing cheating?

Although 85% of the conversations received a high rating for helpfulness, 15% of them received a medium rating. For instance, in conversation 012, when asked to find the derivative of an equation, the assistant simply responded, instructing the student to differentiate each term without explaining how to, such as by mentioning the power rule. This was not very helpful.

To address these issues, future studies should focus on expanding the dataset beyond the 200 conversations used. This was mentioned as a possible limitation in our initial methodology, and the results support it. A larger, more diverse dataset of thousands of conversations would improve the results seen in this model.

However, in comparison to the pre-tuned gpt-3.5-turbo-0125 model, the Intelligent Academic Assistant showed significant improvements. It consistently supported critical thinking by guiding students to the correct answer, not simply providing it.

Overall, the results of the Intelligent Academic Assistant are promising. It distinguishes itself from other models by supporting critical thinking and guiding the student through its prompt. The setbacks encountered in the results can be improved through a larger dataset of

conversations, improving the results of this resource as an effective tool for learning.

9 Discussion and Analysis

9.1 Implications

The Intelligent Academic Assistant successfully met the following two study goals:

- To develop a prototype Generative AI academic assistant that supports critical thinking and learning while mitigating academic dishonesty. This prototype will serve as an alternative to the outright ban of Generative AI in schools.
- To provide a tangible demonstration of a Generative AI resource that supports student learning and cognitive development, thereby encouraging schools to research and implement similar solutions.

Only time will tell whether the third goal of initiating a dialogue will be met, though discussion of ethical GenAI use will begin at the Long Island Youth Summit.

This solution aligned with the thoughts of the educational professionals introduced earlier. Victor Lee, a Stanford education scholar, stated, “AI is not going away. Along with addressing the deeper reasons why students cheat, we need to teach students how to understand and think critically about this technology” [11]. Additionally, Nagelhout [2024] acknowledged that GenAI has benefits beyond cheating, calling it “the modern approach to learning.” This solution demonstrates how GenAI can be used responsibly and ethically by restoring critical thinking and preventing the use of academic dishonesty.

The results of the Intelligent Academic Assistant were astonishing. The model received perfect ratings across all criteria in most conversations. This serves as an alternative to the outright ban of GenAI, which is currently being implemented in schools across Long Island. Just like the introduction of the internet and the calculator into education, GenAI is here to stay, and education must adapt to its use.

9.2 Limitations and Areas for Future Research

The methodology and solution for this project was not perfect and has several areas that can be improved in future studies.

The solution did not fully address the ethical concerns of privacy and security. Although the application has many security features, such as encrypted data, the option to permanently delete conversations, and no training of the model through new conversations, there is always the risk of a data breach that could compromise user data.

The solution did not address the ethical concerns regarding bias. The model used the gpt-3.5-turbo-0125 model which was already biased due to the data used to train it, as explained by Datta [2025]. This issue was beyond our control and was a trade-off due to the complexity of creating a large language model from scratch.

The only results measured were an analysis of responses to certain processes. It was a qualitative analysis, which was subjective, meaning it was influenced by our personal feelings. There was no objective measure of the actual impact on student learning, which did not occur due to the short timeframe for submission to the Long Island Youth Summit. Future studies should include quantitative analysis and measure actual helpfulness.

The most significant limitation was using a small 200 conversation dataset to train the gpt-3.5-turbo-0125 model, which affected the results of the Intelligent Academic Assistant. There is significant room for improvement in balancing direct answers with supporting critical thinking and improving response helpfulness, as seen in the Results section. Additionally, the fine-tuned model did not incorporate an adaptive learning system, which meant that there was no improvement process beyond the initial training. This was done out of privacy concerns in training the model on new conversational data. Future models can benefit from significantly expanding the conversation dataset used to train the model, incorporating a feedback system where users can rate helpfulness, and incorporating adaptive learning for improved results.

Finally, this solution is only targeted toward GenAI use inside of K-12 schools in Long Island, where GenAI usage can be controlled through network and Chromebook protocols. However, outside of school, students can access any resource without limitations. This is a major reason for alternative solutions such as legislation that aims to address this issue.

10 Conclusion

The most popular alternative solution to the one proposed that addresses unethical use of GenAI is legislation on Long Island that bans or limits its use. Schools across Long Island have begun banning GenAI inside of their schools through network and computer protocols that prohibit its access. Additionally, educators have begun catching students through the use of AI detection, such as Turnitin [2024]. These measures are appropriate and warranted and are very successful solutions. Just for clarity, our stance is against GenAI when used for academic dishonesty, as it's considered plagiarism. However, that doesn't mean to ban it. The benefits of AI explained by educational professionals earlier are being overlooked, and legislation banning it would be counterintuitive. Banning AI use across Long Island would put students at a disadvantage and wouldn't be fair to students in other regions, where AI is allowed. The federal government has shown no intention of getting rid of it, in fact investing $500 billion in AI infrastructure [6].

GenAI can be comparable to the internet and the calculator. Legislation ensuring privacy and security is acceptable and needs to be implemented. However, we shouldn't ban its access. Instead, schools must adapt, which has already been taking place through the implementation of AI detection. Cheating has always been an issue, as mentioned by Spector [2023], where surveys over the past 15 years show that 60-70% of students admit to cheating in some form each month. That being said, schools can allow acceptable models similar to our proposed solution to begin transitioning into ethical GenAI use. This proposed solution encourages schools to allow GenAI resources that support student learning while continuing its policies to catch academic dishonesty.

References

[1] ACT. 2023. Half of High School Students Already Use AI Tools. *ACT*. Retrieved from https://leadershipblog.act.org/2023/12/students-ai-research.html

[2] Emily P. Backes and Richard J. Bonnie. 2019. *The Promise of Adolescence: Realizing Opportunity for All Youth*. National Academies Press (US), Washington (DC). Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK545481/

[3] Center for Democracy & Technology. 2024. New Research: Educators Still Struggling with Generative AI Detection, Discipline, and Distrust, Despite Increased School Guidance. *Center for Democracy and Technology*. Retrieved from https://cdt.org/press/new-research-educators-still-struggling-with-generative-ai-detection-discipline-and-distrust-despite-increased-school-guidance/

[4] Anupam Datta. 2025. 3 kinds of bias in AI models -- and how we can address them. *Scoolaid.net*. Retrieved February 17, 2025 from https://link-gale-w9ba.orc.scoolaid.net/apps/doc/A652877175/ITOF?u=nysl_li_amit&sid=bookmark-ITOF&xid=4a9d7564

[5] Michael Gerlich. 2025. AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking. *Societies* 15, 1 (January 2025), 6. DOI:https://doi.org/10.3390/soc15010006

[6] Anirban Ghoshal. 2025. OpenAI, SoftBank, Oracle lead $500B Project Stargate to ramp up AI infra in the US. *Scoolaid.net*. Retrieved February 17, 2025 from https://link-gale-w9ba.orc.scoolaid.net/apps/doc/A824433079/ITOF?u=nysl_li_amit&sid=bookmark-ITOF&xid=944cdb2b

[7] Hank Lee, Advait Sarkar, Lev Tankelevitch, Ian Drosos, Sean Rintel, Richard Banks, and Nicholas Wilson. 2025. The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers. (2025). DOI:https://doi.org/10.1145/3706598.3713778

[8] Ryan Nagelhout. 2024. Students Are Using AI Already. Here's What They Think Adults Should Know. *Harvard Graduate School of Education*. Retrieved from https://www.gse.harvard.edu/ideas/usable-knowledge/24/09/students-are-using-ai-already-heres-what-they-think-adults-should-know

[9] Shane O'Neill. 2025. Ethics of generative AI: To be innovative, you must first be trustworthy. *Scoolaid.net*. Retrieved February 17, 2025 from https://link-gale-w9ba.orc.scoolaid.net/apps/doc/A797735787/ITOF?u=nysl_li_amit&sid=bookmark-ITOF&xid=f3adf34b

[10] Salesforce. 2023. Top Generative AI Statistics for 2023. *Salesforce*. Retrieved from https://www.salesforce.com/news/stories/generative-ai-statistics/#h-General-population-data-reveals-new-insights-

[11] Carrie Spector. 2023. What Do AI Chatbots Really Mean for Students and cheating? *Stanford Graduate School of Education*. Retrieved from https://ed.stanford.edu/news/what-do-ai-chatbots-really-mean-students-and-cheating

[12] The Economist. 2025. A short history of AI. *Scoolaid.net*. Retrieved February 17, 2025 from https://link-gale-w9ba.orc.scoolaid.net/apps/doc/A801498360/ITOF?u=nysl_li_amit&sid=bookmark-ITOF&xid=b55d4869

[13] Tim Mucci. 2024. History of artificial intelligence. *Ibm.com*. Retrieved from https://www.ibm.com/think/topics/history-of-artificial-intelligence

[14] Turnitin. 2024. Turnitin marks one year anniversary of its AI writing detector with millions of papers reviewed globally. *www.turnitin.com*. Retrieved from https://www.turnitin.com/press/press-detail_17793

A Appendix

A.1 200 Conversation Dataset Used to Train the Model

This dataset, composed of 200 conversations, was created to train the GPT-3.5-turbo-0125 model, transforming it into an Intelligent Academic Assistant. The dataset was created by us and is publicly available for free-use at the following GitHub repository: https://github.com/jesbarrios/model-data-set/blob/main/data.jsonl

A.2 Source Code for Intelligent Academic Assistant

The source code for the Intelligent Academic Assistant is open-source and freely available for use. The code can be accessed through the following GitHub repository: https://github.com/jesbarrios/Intelligent-Academic-Assistant

A.3 Intelligent Academic Assistant Testing

ID	Subject	User Input	Response After Fine-Tuning
001	English	“Write a five-paragraph persuasive essay arguing whether social media has a positive or negative impact on modern society. Support your argument with specific examples.”	“Consider starting with a strong thesis that clearly states your position on social media. For example, do you believe it has had an overall negative impact due to issues like cyberbullying and misinformation, or do you see it as a positive tool for connecting people from all over the world? How can you support this thesis with strong, well-researched examples that show the broader implications of social media use?’
002	English	“Examine the character development of	“In examining Elizabeth Bennet's character, think about her initial judgments and attitudes, especially towards Mr. Darcy. What key events challenge these beliefs and lead to her personal growth? How do her interactions with other characters,

		Elizabeth Bennet in 'Pride and Prejudice.' How does she change throughout the novel, and what are the key events that contribute to her transformation?"	such as Jane or Lady Catherine, shape her perspective and influence her decisions? Try linking these events to the underlying themes of prejudice and social class in the novel. How does Austen use Elizabeth's transformation to address these issues?"
003	English	"Explain the difference between dramatic irony, situational irony, and verbal irony."	"Dramatic irony occurs when the audience or reader knows something that the characters do not, often used to build tension. Situational irony happens when the outcome of a situation is different from what is expected, sometimes with a humorous twist. Verbal irony is when someone says something but means the opposite, often used to express sarcasm. Can you think of examples of each type of irony from a favorite book or movie? How do they add depth to the story or create a comedic effect?"
004	English	"Write an essay discussing the role of setting in creating atmosphere in Edgar Allan Poe's short stories."	"Focus on how Poe's gothic settings contribute to the dark, eerie mood of his tales. For example, what impact does a decaying mansion or a stormy night have on the reader's emotions and expectations? How does the setting influence the characters' actions and decisions? Think about why Poe's use of setting is crucial to building suspense and how it reflects the themes of his work. How does the setting help the stories stand out in the horror genre?"
005	English	"Compare the differing uses of first person narrative in 'To Kill a Mockingbird' and 'The Great Gatsby'."	"Consider the different voices and perspectives created by Scout Finch in "To Kill a Mockingbird" and Nick Carraway in "The Great Gatsby." How does each narrator's viewpoint affect the reader's understanding of the story? Analyze how the first-person narratives shape the reader's connection to the characters. What unique insights do Scout and Nick provide because of their personal involvement in the events? How does the use of first-person enhance the themes of morality, truth, and perception in each novel?"
006	Social Studies	"Summarize the main causes of the American Civil War, and list them in order of importance."	The causes of the American Civil War are multifaceted, ranging from economic differences to the moral issue of slavery. Key factors include debates over states' rights, economic disparities between the North and South, and, significantly, the moral clash of slavery. Organize the causes by impact rather than importance—reflect on how the combination of these factors ultimately led to conflict. How did each cause contribute to the escalating tensions that resulted in war?
007	Social Studies	"Compare and contrast the political systems of ancient Athens and ancient Rome."	"Consider how both city-states' political systems were rooted in democracy but had key differences in terms of citizen participation and decision-making. Highlight Athens' direct democracy, where citizens voted on laws, and Rome's representative Republic, where elected officials made decisions. Discuss how these systems reflected the values of each society. How did the balance of power between the people and the ruling elite shape the course of history in each city-state?"
008	Social Studies	"Explain the impact of the Industrial	"Explore how the Industrial Revolution altered the traditional class structure. You might want to discuss the rapid growth of the working class, the emergence of a middle class, and the continued power of the aristocracy. How did industrialization

		Revolution on social classes in 19th-century England."	contribute to the widening gap between the wealthy and the impoverished? Think about how figures like Karl Marx addressed these social changes and why this period is often associated with class warfare."
009	Social Studies	"Create a timeline of the major events of World War II, including dates and brief descriptions."	"World War II was a complex conflict that lasted from 1939 to 1945. Start with the invasion of Poland in September 1939, which was a turning point that led to the involvement of major powers. Then, move on to key battles like the Battle of Stalingrad, which occurred from 1942 to 1943 and was a turning point on the Eastern Front. The D-Day invasion of Normandy in June 1944 marked the beginning of the end for Germany, leading to their surrender in May 1945. Don't forget to include significant events like the bombing of Hiroshima and Nagasaki in August 1945, which ultimately led to Japan's surrender and the end of the war. Try to make connections between events to show how they influenced each other."
010	Social Studies	"Define the concept of 'checks and balances' in the U.S. government, and give an example of how it works."	"Checks and balances are a fundamental principle of the U.S. government that ensures no branch becomes too powerful. For example, the President can veto legislation passed by Congress, but Congress can override that veto with a two-thirds majority vote. Additionally, the Supreme Court can declare acts of Congress unconstitutional, acting as a check on the legislative branch. You could also discuss how the Senate confirms presidential appointments, providing a check on the executive branch's authority in selecting key officials. How do you think this system prevents tyranny and promotes accountability?"
011	Math	"Solve this quadratic equation: $2x^2 + 5x - 3 = 0$, and show all the steps."	"This is a quadratic equation, and you can solve it using the quadratic formula: $x = [-b \pm \sqrt{(b^2 - 4ac)}] / 2a$. For this equation, $a = 2$, $b = 5$, and $c = -3$. First, calculate the discriminant, $b^2 - 4ac$, then substitute the values into the quadratic formula to find the solutions. How would you rearrange the equation and solve for the roots?"
012	Math	"Calculate the derivative of the function $f(x) = x^3 - 4x^2 + 7x - 2$."	"To calculate the derivative of a polynomial, you can differentiate term by term. What do you get when you find the derivatives of each term in this function?"
013	Math	"Explain how to find the area of a regular hexagon with a side length of 8 cm."	"For a regular hexagon, you can think of it as six equilateral triangles. The formula for the area of an equilateral triangle is ½ base × height. The base would be the side length of the hexagon, which is 8 cm. To find the height, you can draw a line from one vertex, create a right triangle, and use the Pythagorean theorem. Can you determine the area of an equilateral triangle and then the full hexagon based on that information?"
014	Math	"Find the slope of the line that passes through the points (2, 5) and (6, 13)."	"To find the slope of a line, you can use the formula $m = (y_2 - y_1) / (x_2 - x_1)$. Can you substitute the values from the points into this formula?"
015	Math	"A right triangle has legs with lengths of 6 cm and 8 cm. What is the length of the hypotenuse?"	"This problem seems to involve the Pythagorean Theorem! It states that in a right triangle, the square of the length of the hypotenuse is equal to the sum of the squares of the other two sides. Can you apply this theorem with the given lengths?"
016	Science	"Explain the	"Photosynthesis is a vital process in plants, where they convert light energy into

		process of photosynthesis, including the chemical equation."	chemical energy in the form of glucose. The chemical equation for photosynthesis is $6CO_2 + 6H_2O$ + light energy → $C_6H_{12}O_6 + 6O_2$. This equation shows that carbon dioxide and water, in the presence of light, are used by plants to produce glucose and release oxygen. Can you discuss how light energy is converted into chemical energy in this process?"
017	Science	"Describe the structure of a DNA molecule and explain its role in heredity."	"DNA is a double-helix molecule made up of two long strands tightly wound around each other. Each strand contains sequences of nucleotides, which are the building blocks of DNA. The nucleotides are composed of a sugar, a phosphate group, and one of four bases: adenine (A), thymine (T), cytosine (C), or guanine (G). The bases form pairs across the two strands: A pairs with T, and C pairs with G. This structure allows DNA to carry genetic information through generations. During cell division, the DNA is replicated, ensuring that each new cell receives the same genetic material. The order of the bases along the DNA strand forms a genetic code that determines an organism's traits, such as eye color or susceptibility to certain diseases. How do you think variations in this genetic code lead to the diversity of traits among different organisms?"
018	Science	"Calculate the acceleration of an object that goes from rest to 25 m/s in 5 seconds."	"To calculate acceleration, you use the formula a = (vf - vi) / t, where vf is the final velocity, vi is the initial velocity, and t is the time. In this case, the initial velocity (vi) is 0 m/s, the final velocity (vf) is 25 m/s, and the time (t) is 5 seconds. Can you substitute these values into the formula and calculate the acceleration?"
019	Science	"List the main layers of the Earth's atmosphere, and give a brief description of each."	"The Earth's atmosphere has layers, including the troposphere, stratosphere, mesosphere, thermosphere, and exosphere. The troposphere is the closest to the Earth's surface and where weather occurs. The stratosphere contains the ozone layer, which absorbs most of the sun's harmful ultraviolet radiation. The mesosphere is where meteors burn up upon entry into the atmosphere. The thermosphere is where auroras occur, and the exosphere is the outermost layer that fades into space. How do you think these layers protect Earth and living organisms?"
020	Science	"Explain the difference between mitosis and meiosis, and describe the purpose of each process."	"Mitosis and meiosis are both processes of cell division but serve different purposes. Mitosis occurs in somatic (body) cells and is responsible for growth, repair, and asexual reproduction. It produces two identical daughter cells, each with the same number of chromosomes as the parent cell. On the other hand, meiosis occurs in reproductive cells and is essential for sexual reproduction. It involves two rounds of cell division, resulting in four cells with half the number of chromosomes as the parent cell. This leads to genetic diversity in offspring by shuffling and combining genetic information from two parents. Can you think of an example of when meiosis plays a crucial role in creating genetic variation?"